PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	ABCD Criteria to Improve Visual Inspection with Acetic Acid (VIA) Triage in HPV-positive Women: a Prospective Study of Diagnostic Accuracy
AUTHORS	Petignat, Patrick; Kenfack, Bruno; Wisniak, Ania; Saiji, Essia; Tille, Jean-Christophe; Tsuala Fouogue, Jovanny; Catarino, Rosa; Tincho, Eveline; Vassilakos, Pierre

VERSION 1 – REVIEW

REVIEWER	Schiffman, Mark National Cancer Institute, NIH, DHHS
REVIEW RETURNED	12-Jul-2021

GENERAL COMMENTS	Review of "ABCD criteria..." Reviewer: Mark Schiffman, MD, MPH, NCI, NIH Maryland, USA This manuscript reports on the performance of the authors' ABCD visual classification system for the management of women found to be HPV-positive in low-resource settings. They describe their experience in Cameroon. The field study was delayed and shortened due to COVID-19. However, it remains a serious and thoughtful effort to standardize visual evaluation, which tends to be poorly reproducible and inaccurate unless very careful quality assurance is maintained. The use of VIA for triage of HPV-positives increases difficulty. Even experts have difficult distinguishing between minor changes of HPV infection and normal look-alike changes, and between HPV and precancerous lesions at the borderline, or equivocal "gray zone" of uncertainty between the two (Massad et al., JLGTD 2009). Classification schema like the Reid Index and the Swede Score combine several features to produce an overall score that is divided to categorize the overall colposcopic severity. The authors here present use of an ABCD index. I have the following suggestions: 1. Regardless of past literature, it is important to explain ABCD more fully. The statement that it represents past work, with

	references, was not enough for me to know why the features were chosen. Acetowhitening and bleeding are highlighted. I was hoping to see details of the weighting of the features, and how that weighting was validated. We also published that acetowhitening is sensitive but not specific. The data is the paper show that sensitivity for precancer (CIN2, or CIN3) can be achieved by calling positive all acetowhite lesions, but the specificity falls below 50%. Therefore, VIA is calling positive about 60% of women with HPV positivity. In a setting with 15-30% HPV positivity, the numbers of women that will be treated can rise to 15%. If ablation is not possible, then a large number of referrals/excisions will be performed due to HPV/VIA positivity, although we have insufficient prospective data regarding the risk of precancer in thisgroup. In our work in Nigeria, we find fewer cases of precancer among HPV-positive women than we expected, and we remain uncertain as to the probability of precancer given HPV. In a paper by Campos et al. (2021) in Preventive Medicine, we call for more research into whether the natural history of HPV is different in high prevalence regions, calling for caution in universal treatment particularly excision. In short, if the PPV is low using the strategy in the paper, is excision justified only if very safe and performed by experts. Are these numbers of referrals practical? 2. The paper seemed to represent the conclusions of the authors group of how VIA in Cameroon and similar places could be standardized. In the US management guidelines, we separated risk estimation from clinical action decisions. A group of clinicians led consideration of who to treat for what level of risk, given US risk tolerance and resources. In addition to wanting more details of the evidence for the accuracy and reliability risk estimations (i.e., the weighting of the features in ABCD), I was hoping to see more about clinical decision making as to how the ABCD scores and HPV data would be integrated by local clinicians to decide who to treat and how. The ABCD performance of good sensitivity but very high referral rates left open important questions, and it was not clear that the scale provides a sustainable screening program result in settings where treatment capacity is low. This is important ongoing work.

REVIEWER	Cubie, Heather The University of Edinburgh
REVIEW RETURNED	29-Sep-2021

GENERAL COMMENTS	This is a very clearly and authoratatively written manuscript describing the use of a simple and memorable set of criteria (ABCD) to increase the accuracy oof triage of HR-HPV+ women undergoing cervical screening in LMIC. The analysis is based on 340 results which had HPV, ABCD and histology results. The cytology performed in Geneva adds little to the manuscript. The sensitivity of ABCD triage was greater than for HPV alone but at a cost of lower specificity and thought should be given to whether the criteria could be further refined to improve specificity. For example, although not presented, would there have been any improvement had partial genotyping (HPV 16, 18/45 or other) been included (ABCDG)?

	The ABCD criteria proved useful and independent of the multiple sociodemographic characteristics which were analysed. Perhaps it is time to reduce the publication space of such demographic details which in many studies have been found to be non-impactful. This study was carried out in a single district in West Cameroon and the authors acknowledge that assessment across other locations is needed to make the findings generalisable. The discussion spends some time on the overtreatment even with ABCD criteria, but considers this might be a price worth paying in LMIC for same day services, reduced loss to follow-up, low morbidity of thermal ablation and indeed, the confidence of women in a comprehensive service. In my view, these considerations cannot be made strongly enough.

REVIEWER	Baena, Armando International Agency for Research on Cancer
REVIEW RETURNED	01-Oct-2021

GENERAL COMMENTS	This is a very interesting study on the triage of HPV positive women in a sub-Saharan African country using enhance VIA. Despite the limited sample size, results are relevant, especially in light of the launch of the new WHO cervical cancer screening guidelines and the global cervical cancer elimination initiative. Some comments for the authors are presented below. Introduction I suggest authors highlight the relevance of the topic mentioning the WHO cervical cancer elimination initiative, particularly the screening component and how important alternative/enhance screening methods such as VIA may contribute to this initiative. Some references should be updated; for instance, Globocan 2020 (instead of 2018), WHO cervical cancer screening and treatment 2021 (instead of 2013). Page 7, lines 6 and 7: please revise the definition of a valid triage system/test and elaborate more this concept/idea. Authors mention that a valid triage system is the one that conserves the high sensitivity of the HPV test for the detection of high-grade cervical lesions which is partially true. Authors may have missing the role of the positive predictive value (PPV) of a triage test (the higher the better) and the referral rate. Even triage tests with limited sensitivity but with great PPV and referral rate are desirable especially when the lost to follow-up of women referred to surveillance is minimised. Reference number 10 (IARC VIA manual) seems to be swapped, please check. Page 8, line 13: please briefly explain/expand the ABCD criteria here. Methods Please clarify if it was also the primary VIA examiner in charge of evaluating the transformation zone. From the results, it seems that a TZ type 3 was evaluated with the ABCD criteria and the VIA was not considered “not evaluable”. Was the TZ an independent variable of the ABCD result? ECC is described in the histology findings section. Please clarify how the ECC was processed (cellular block or smear). If it was a smear, was it part of the outcome? If so, please include in the discussion the implications of using an ECC smear for the outcome. Was the histology externally reviewed? If so, please include a statement explaining the review process. If not, please clarify it. Page 12, line 12: “Patients were” instead of “Patients are”?

	Results Table 2: given that the outcome was define as a concordance between a positive ABCD result and CIN2+ based on histology, including CIN2+ based on histology as an explanatory variable is not a valid approach and this variable should be removed from the analyses. The analysis of the enhance VIA as a triage test misses some concepts that are relevant in the context of the evaluation of a triage test of HPV positive women. For instance, the referral rate. Although it is indirectly presented in Table 1 (percentage of ABCD positivity, i.e., 60.9%), this important performance measure is not explicitly presented or discussed as part of the reasons of the high Sensitivity and low Specificity. It is not either compared to other triage tests. I suggest authors include a column in Table 3. Other concept is the interpretation of the PPV and NPV in terms of CIN2+ risk and clinical management. The PPV of VIA is very limited compared to other triage tests (HPV16/18 genotyping for instance), and below of 20% for CIN2+ which is a known referral threshold for CIN2+. Discussion I suggest authors discuss differences between VIA/VILI alone vs enhance VIA/VILI. Maybe it would be worth including some results with these comparisons to evaluate the value of adding digital images and external clinicians' support. Authors mentioned that the generalisability of the study is good given the lack of associations of outcomes with socio-demographic variables (page 21). However, because of the sample size and limited study centres included, this comment should be reconsidered and/or elaborated more appropriately. Strengths of the study are not discussed, and limitations are not fully covered (e.g.: sample size) and those that were mentioned are not widely explained; for instance, implications of having included a single centre. Implications of enhancing VIA with digital images should be more elaborated. Also, the feasibility of implementing this approach. Authors mention their 3T-Approach in the discussion, but it was barely mentioned in Methods.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Mark Schiffman, National Cancer Institute, NIH, DHHS Comments to the Author:

This manuscript reports on the performance of the authors' ABCD visual classification system for the management of women found to be HPV-positive in low-resource settings. They describe their experience in Cameroon. The field study was delayed and shortened due to COVID-19. However, it remains a serious and thoughtful effort to standardize visual evaluation, which tends to be poorly reproducible and inaccurate unless very careful quality assurance is maintained. The use of VIA for triage of HPV-positives increases difficulty. Even experts have difficult distinguishing between minor changes of HPV infection and normal look-alike changes, and between HPV and precancerous lesions at the borderline, or equivocal "gray zone" of uncertainty between the two (Massad et al., JLGTD 2009). Classification schema like the Reid Index and the Swede Score combine several

features to produce an overall score that is divided to categorize the overall colposcopic severity. The authors here present use of an ABCD index.
I have the following suggestions:

1. Regardless of past literature, it is important to explain ABCD more fully. The statement that it represents past work, with references, was not enough for me to know why the features were chosen. Acetowhitening and bleeding are highlighted. I was hoping to see details of the weighting of the features, and how that weighting was validated.

Thank you for this remark and we can add the following observations to explain how these criteria have been developed and implemented in our setting.

Criterion A (acetowhiteness): Guidelines (as those defined by the IARC or IFCPC) consider that acetowitheness is of important diagnostic value for high-grade CIN diagnosis if it is a “dense acetowhitening area”. The “denseness of acetowhiteness” has been explored in different studies, and it appears that subtle signs like “thin or mild or shady or translucent acetowhite change” may also be associated with presence of CIN2+ lesions. In the Swede score too, “the degree of acetowhiteness” has been evaluated and it has been reported that the use of a low cut-off score (corresponding to a low degree of acetowhiteness) allows to reach a very high sensitivity (100% (95%CI, 89.6%-100%)) for the detection of high-grade lesions (Ranga et al. 2017; Strander et al. 2015). Therefore, in order to optimize the sensitivity of the test, we have considered here any acetowhitening of 5 mm or larger as ABCD-positive.
Criteria B (bleeding): Presence of bleeding (small spots in the TZ area) is easily recognized by front-line health care providers. Naked-eye evaluation enhanced by camera or smartphone-acquired digital imaging does not reach the quality of image of high-resolution colposcopy, nor does it allow to observe all the small signs that can be seen in a true colposcopy exam. However, presence of cervical bleeding in the TZ area is generally easily recognized. In a study (Basu et al. 2002) conducted in India with a “screen-and-treat” approach, criteria like “bleeding on touch” or “bleeding erosion” were considered as a “high-threshold positive” for cervical cancer diagnosis. The weakness of this sign is that inflammation or infection can also be associated with ulcers and cervical bleeding in the TZ. Currently, our data doesn’t allow to specifically determine the sensitivity and specificity of this specific sign, but we have an on-going trial which should give us more information about the performance of criterion B in a screening context.

These issues have been restructured in the methods section, lines 124-152:

- Criterion A for Acetowhiteness – Criterion A is obtained after application of 3%–5% acetic acid. Any acetowhite area touching the TZ and having a diameter of >5 mm (criterion D) is considered positive. Compared with the IARC criteria, which require a degree of whiteness together with the presence of a sharp, distinct, well defined, dense (opaque/dull or oyster white) acetowhite area,(12) we considered here any acetowhite lesion exceeding 5 mm to be positive.
- Criterion B for Bleeding on touch – Criterion B is obtained upon native examination or after acetic acid application. Presence of cervical bleeding without touching or after lightly touching the cervix in the TZ area is considered positive. This means that any bleeding from the surface of the cervix, after excluding bleeding of intra-uterine origin, can be associated with CIN2+ lesions. Although bleeding can also be caused by ulceration or infection, any signs should be thoroughly investigated to rule out the possibility of early preclinical invasive cancer. This sign is easy to recognize and is considered a high-risk finding for precancerous lesions and cervical cancer.(24,25) Presence of bleeding in association with criteria A and C may require further testing like biopsy or loop electrosurgical excision.

- Criterion C for Colouring with Lugol's iodine – Criterion C is optional. Lugol's iodine staining can be used as an adjunct to VIA to recognize epithelial change that would otherwise be difficult to identify by VIA only. The colour changes with VILI can be easier to appreciate than those after VIA and may contribute to identification of a missed thin acetowhite lesion. To be considered positive, an iodine-negative lesion should correspond to a VIA lesion having criteria A and D. Compared with the IARC criteria, which require the presence of a well-defined, bright yellow, iodine non-uptake area,(12) we consider any non-iodine uptake areas to be positive, providing they match an acetowhite lesion.
- Criterion D for Diameter – Criterion D is evaluated after application of acetic acid (or Lugol's iodine). An acetowhite lesion measuring >5 mm in diameter (about the size of a pencil eraser) is considered positive. Defining a minimal size of 5 mm allows exclusion of benign conditions such as dot-like, line-like, or streak-like areas.(23)

2. We also published that acetowhitening is sensitive but not specific. The data is the paper show that sensitivity for precancer (CIN2, or CIN3) can be achieved by calling positive all acetowhite lesions, but the specificity falls below 50%. Therefore, VIA is calling positive about 60% of women with HPV positivity. In a setting with 15-30% HPV positivity, the numbers of women that will be treated can rise to 15%. If ablation is not possible, then a large number of referrals/excisions will be performed due to HPV/VIA positivity, although we have insufficient prospective data regarding the risk of precancer in this group. In our work in Nigeria, we find fewer cases of precancer among HPV-positive women than we expected, and we remain uncertain as to the probability of precancer given HPV. In a paper by Campos et al. (2021) in Preventive Medicine, we call for more research into whether the natural history of HPV is different in high prevalence regions, calling for caution in universal treatment particularly excision. In short, if the PPV is low using the strategy in the paper, is excision justified only if very safe and performed by experts. Are these numbers of referrals practical?

Thank you for this comment. We fully agree that acetowhitening is sensitive but not specific. In our setting, we considered that the priority was to have a highly sensitive triaging method. In our previous experiences (Tebeu et al., IJC 2015; Bigoni et al., IJC 2015) using "traditional VIA criteria" (dense acetowhite, dense opaque grey-white areas), we found that the gain in specificity when adding VIA to HPV testing was obtained at the expense of an important loss in sensitivity, which we considered unacceptable in thcontext of a mass screening program. Therefore, the current approach was to lower the VIA positivity threshold in order to gain in sensitivity.
Our approach, as compared to the WHO option "HPV test and treat", offers a significant gain in specificity.
We also agree that the low PPV limits this strategy if participants having a positive screen need to be referred for further investigation (colposcopy, biopsy/histology, LLETZ), especially in a low-resource setting, where referral services are not readily available. This has been addressed in the manuscript.

Lines 334-353: Despite the low specificity, our 3T-Approach in a single visit may be acceptable in an LMIC context because it reduces cost and loss to follow-up, which are recognized barriers to effective cervical cancer screening.(11,27) Indeed, studies in Uganda(28) and South Africa(27) have shown loss to follow-up rates between 21% and 25% after the first visit, up to 50% at 24 months. Furthermore, treatment by thermal ablation is associated with very low risks of side effects and morbidity.(29) Therefore, treatment of a significant number of false-positive cases may be considered an acceptable strategy for effective control of CC in an LMIC setting and may contribute to reaching the target of the WHO's elimination initiative.(3,5) However, the use and integration of the ABCD criteria in the cervical cancer screening process warrants multidisciplinary discussion with involved stakeholders, taking into account the local context and resources, as well as regional HPV prevalence, prevalence of CIN2+ in HPV-positive participants, level of risk including HIV prevalence, availability of treatment modalities on site, and the possibility to offer further investigation when

required. According to the context, the decision to refer has consequences for the patients and the health care system, requiring additional time and resources, and increasing the risk of loss to follow-up. Recognizing the limitations of the ABCD criteria with regard to PPV and overtreatment rates, other triaging strategies merit further investigation. The use of extended HPV genotyping (HPV 16, 18, 45, 31, 33, 35, 52 and/or 58) for the triaging of HPV-positive women is one alternative that should also be explored.

3. The paper seemed to represent the conclusions of the authors group of how VIA in Cameroon and similar places could be standardized. In the US management guidelines, we separated risk estimation from clinical action decisions. A group of clinicians led consideration of who to treat for what level of risk, given US risk tolerance and resources. In addition to wanting more details of the evidence for the accuracy and reliability risk estimations (i.e., the weighting of the features in ABCD), I was hoping to see more about clinical decision making as to how the ABCD scores and HPV data would be integrated by local clinicians to decide who to treat and how.

Thank you for this valuable comment.
The WHO algorithm for primary HPV screening strategies includes (i) HPV and treat, or (ii) HPV, triage and treat. In the area under study and according to previous studies, we expected a prevalence of HPV of approximately 15% and a prevalence of CIN2+ in HPV-positive women of 10-15%. As more than 80% of women having an HPV-positive test do not have any CIN2+, the involved partners decided that the most valuable approach was to add triage to improve specificity. The discussion has been completed according to your recommendations.

Lines 342-349: However, the use and integration of the ABCD criteria in the cervical cancer screening process warrants multidisciplinary discussion with involved stakeholders, taking into account the local context and resources, as well as regional HPV prevalence, prevalence of CIN2+ in HPV-positive participants, level of risk including HIV prevalence, availability of treatment modalities on site, and the possibility to offer further investigation when required. According to the context, the decision to refer has consequences for the patients and the health care system, requiring additional time and resources, and increasing the risk of loss to follow-up.

The ABCD performance of good sensitivity but very high referral rates left open important questions, and it was not clear that the scale provides a sustainable screening program result in settings where treatment capacity is low.

Please refer to the response to comment #2 above, where we have addressed the limitations of the strategy in the discussion. Of note, ABCD-positive women do not require referral nor an additional pelvic exam, as they are immediately treated at the end of the visual assessment procedure.

Reviewer: 2
Prof. Heather Cubie, The University of Edinburgh Comments to the Author:

This is a very clearly and authoratatively written manuscript describing the use of a simple and memorable set of criteria (ABCD) to increase the accuracy of triage of HR-HPV+ women undergoing cervical screening in LMIC.

The analysis is based on 340 results which had HPV, ABCD and histology results. The cytology performed in Geneva adds little to the manuscript.

Thank you for this comment. However, we believe that including the performance of cytology as a triage strategy allows to compare the ABCD criteria with a standard screening procedure used in most high-income countries.

The sensitivity of ABCD triage was greater than for HPV alone but at a cost of lower specificity and thought should be given to whether the criteria could be further refined to improve specificity. For example, although not presented, would there have been any improvement had partial genotyping (HPV 16, 18/45 or other) been included (ABCDG)?
We agree with this comment, and there probably remains scope for improvement in the triage of HPV-positive women. We have addressed this issue by evaluating the performance of partial genotyping (HPV 16, 18 and/or 45) as a second triage strategy for ABCD-positive patients, which indeed improves specificity, however at the cost of a large loss in sensitivity. Using our available data, specificity for this combined triaging strategy was estimated at 88.3%, with a sensitivity of 30%. While these values may be acceptable in settings where patient follow-up is high, it is not suited to a context with high risk of loss to follow-up. We are currently evaluating the use of extended genotyping as an alternative triaging strategy, which we have added to the discussion.

Lines 349-353: Recognizing the limitations of the ABCD criteria with regard to PPV and overtreatment rates, other triaging strategies merit further investigation. The use of extended HPV genotyping (HPV 16, 18, 45, 31, 33, 35, 52 and/or 58) for the triaging of HPV-positive women is one alternative that should also be explored.

The ABCD criteria proved useful and independent of the multiple sociodemographic characteristics which were analysed. Perhaps it is time to reduce the publication space of such demographic details which in many studies have been found to be non-impactful.
Thank you for this comment, but we believe that the lack of association of ABCD criteria performance with sociodemographic characteristics supports the generalizability of our results to a wide range of women. Therefore, we would prefer to keep these results in the paper.

This study was carried out in a single district in West Cameroon and the authors acknowledge that assessment across other locations is needed to make the findings generalisable. The discussion spends some time on the overtreatment even with ABCD criteria, but considers this might be a price worth paying in LMIC for same day services, reduced loss to follow-up, low morbidity of thermal ablation and indeed, the confidence of women in a comprehensive service. In my view, these considerations cannot be made strongly enough.
Thank you for this comment. We have strengthened our arguments in the paragraph below.

Lines 334-342: Despite the low specificity, our 3T-Approach in a single visit may be acceptable in an LMIC context because it reduces cost and loss to follow-up, which are recognized barriers to effective cervical cancer screening in Sub-Saharan Africa.[11,27] Indeed, studies in Uganda[28] and South Africa[27] have shown loss to follow-up rates between 21% and 25% after the first visit, up to 50% at 24 months. Furthermore, treatment by thermal ablation is associated with very low risks of side effects and morbidity.[29] Therefore, treatment of a significant number of false-positive cases may be considered an acceptable strategy for effective control of CC in an LMIC setting and may contribute to reaching the target of the WHO's elimination initiative.

Reviewer: 3
Dr. Armando Baena, International Agency for Research on Cancer Comments to the Author:

This is a very interesting study on the triage of HPV positive women in a sub-Saharan African country using enhance VIA. Despite the limited sample size, results are relevant, especially in light of the launch of the new WHO cervical cancer screening guidelines and the global cervical cancer elimination initiative. Some comments for the authors are presented below.

Introduction
I suggest authors highlight the relevance of the topic mentioning the WHO cervical cancer elimination initiative, particularly the screening component and how important alternative/enhance screening methods such as VIA may contribute to this initiative.
Thank you for this suggestion. We have completed the introduction as follows:

Lines 51-56: A global strategy for the elimination of cervical cancer has been launched by the World Health Organization (WHO) in 2020, which relies upon the screening of 70% of women using a high-performance test and the treatment of 90% of women identified with cervical disease.[5] Recommendations adopted by the WHO for screening in resource-limited settings include a strategy of HPV-screening followed by VIA triage and treatment, or a strategy of HPV-screening followed by treatment.[3]

Some references should be updated; for instance, Globocan 2020 (instead of 2018), WHO cervical cancer screening and treatment 2021 (instead of 2013).
The references have been updated as recommended.

Page 7, lines 6 and 7: please revise the definition of a valid triage system/test and elaborate more this concept/idea. Authors mention that a valid triage system is the one that conserves the high sensitivity of the HPV test for the detection of high-grade cervical lesions which is partially true. Authors may have missing the role of the positive predictive value (PPV) of a triage test (the higher the better) and the referral rate. Even triage tests with limited sensitivity but with great PPV and referral rate are desirable especially when the lost to follow-up of women referred to surveillance is minimised.
Thank you for this comment. The introduction has been completed according to your suggestions.

Lines 63-70: A triage system is only a valid option if it can improve the positive predictive value (PPV) for CIN2+ and minimize the referral rate, while conserving the high sensitivity of the HPV test. The achievement of a high PPV at the cost of limited sensitivity may be considered a reasonable option when the loss to follow-up of women requiring surveillance is minimal. However, in low-resource settings, high levels of loss to follow-up constitute an important barrier to cervical cancer screening, which is why programs having no follow-up visits or as few as possible are preferable to achieve a high degree of participation.[11]

Reference number 10 (IARC VIA manual) seems to be swapped, please check.
Thank you for highlighting this. The references have been corrected.

Page 8, line 13: please briefly explain/expand the ABCD criteria here.
The last paragraph of the introduction has been modified as follows:

Lines 88-91: To improve VIA/D-VIA interpretation as a triage test in HPV-positive populations, we introduced a set of criteria, termed ABCD criteria for “Acetowhiteness”, “Bleeding”, “Colouring” (with Lugol’s iodine) and “Diameter” of the lesion. These criteria constitute a simple structure that may contribute to preventing CC in an LMIC context.

Methods
Please clarify if it was also the primary VIA examiner in charge of evaluating the transformation zone.
Yes, the front-line health care providers evaluated the TZ and determined the type of treatment.

Lines 110-112: Healthcare providers performed gynecologic examination with VIA/VILI, assessment of ABCD criteria and transformation zone (TZ) type, and determined treatment modalities in a single visit.

From the results, it seems that a TZ type 3 was evaluated with the ABCD criteria and the VIA was not considered “not evaluable”. Was the TZ an independent variable of the ABCD result?

Eligibility for treatment in the 3T Study was decided independently from TZ interpretation. This means that patients harboring a TZ3 and having no lesion after VIA assessment were considered negative. Only women having a lesion which was not completely visible or extended into the endocervical canal, or a lesion which could not be covered by the probe (e.g. distorted cervix) or suspicious of invasive cancer were considered as ineligible for TA and referred for further evaluation.
In patients considered eligible for TA, we used a narrow probe that reaches the area around the cervical os (initial portion of the endocervical canal) at the end of the procedure to ensure that the entire TZ was treated and to optimize the extent of ablation.
As all of our participants had routine ECC assessment, it was possible to determine prevalence disease in this subgroup of participants (TZ3 and VIA negative). Thus, in our sample, 40 women harbored a TZ3 and 31 of those were considered VIA-negative, among which only one CIN2 was revealed on histology (3.2% of VIA-negative TZ3).

ECC is described in the histology findings section. Please clarify how the ECC was processed (cellular block or smear). If it was a smear, was it part of the outcome? If so, please include in the discussion the implications of using an ECC smear for the outcome.
Was the histology externally reviewed? If so, please include a statement explaining the review process. If not, please clarify it.
The section on histologic analysis has been clarified:

Lines 179-182: Biopsy slides and ECC samples (processed by cellular block) were read by two experienced gynaecologic pathologists of the Geneva University Hospitals, Switzerland, who were blinded to the screening test results and ABCD criteria findings. There was no external review of histological analyses.

Page 12, line 12: “Patients were” instead of “Patients are”?
Thank you for highlighting this. The sentence in the “Patient and public involvement” section has been modified.

Lines 192-193: Patients were also involved at their arrival at the screening center where they were offered a one-hour information session on cervical cancer and sexual health by trained midwives.

Results
Table 2: given that the outcome was define as a concordance between a positive ABCD result and CIN2+ based on histology, including CIN2+ based on histology as an explanatory variable is not a valid approach and this variable should be removed from the analyses.

Thank you for this comment. The variable “CIN2+” has been removed from the analysis (Table 2).

The analysis of the enhance VIA as a triage test misses some concepts that are relevant in the context of the evaluation of a triage test of HPV positive women. For instance, the referral rate. Although it is indirectly presented in Table 1 (percentage of ABCD positivity, i.e., 60.9%), this

important performance measure is not explicitly presented or discussed as part of the reasons of the high Sensitivity and low Specificity. It is not either compared to other triage tests. I suggest authors include a column in Table 3.

A column with the positivity rate of each screening/triage strategy has been added to Table 3. We did not name this 'referral rate' as suggested, because in the case of VIA, patients are treated the same day, during the same pelvic examination and by the same health care provider as the triage test, therefore we cannot call this a referral in the strict sense.

Other concept is the interpretation of the PPV and NPV in terms of CIN2+ risk and clinical management. The PPV of VIA is very limited compared to other triage tests (HPV16/18 genotyping for instance), and below of 20% for CIN2+ which is a known referral threshold for CIN2+.

This issue was addressed in the discussion:

Lines 326-334: Compared to screening by HPV-16/18/45 genotyping without triage, the sensitivity of the ABCD criteria was much higher, at the cost of a lower specificity. PPV was also slightly lower with triage by ABCD criteria (15·1%) than with HPV genotyping (20·9%). Overall, 54·4% of normal histology results and 71·4% of CIN1 were considered ABCD criteria positive and consequently underwent unnecessary treatment. Thus, 85% (174 of 205) of women who screened positive were treated unnecessarily. However, when considering all women screened for CC, including HPV-negative, 174 were treated unnecessarily out of 1964 screened by Self-HPV, corresponding to an overall 8·9% overtreatment rate in the total population screened.

We would also like to emphasize that primary HPV testing followed by triage with VIA is a screening strategy recommended by the WHO for low-resource settings despite the relatively high overtreatment rates. Of note, the alternatively recommended strategy by the WHO of treatment of all HPV-positive women without triage leads to even higher overtreatment rates, which is the basis of the rationale to implement triaging with VIA.

Discussion
I suggest authors discuss differences between VIA/VILI alone vs enhance VIA/VILI. Maybe it would be worth including some results with these comparisons to evaluate the value of adding digital images and external clinicians' support.

We agree that the added value of digital images is an important topic, which we are currently investigating in a sub-sample of our study population. However, in the sample considered in the present analysis, participants did not have a separate diagnosis through VIA and D-VIA. Therefore, the ABCD results presented here are the combined diagnosis of VIA and D-VIA.

Authors mentioned that the generalisability of the study is good given the lack of associations of outcomes with socio-demographic variables (page 21). However, because of the sample size and limited study centres included, this comment should be reconsidered and/or elaborated more appropriately.

Thank you for this comment. We have completed the discussion around generalizability in the following paragraph:

Lines 320-325: The lack of association between multiple socio-demographic variables and a correct prediction of the ACBD criteria (Table 2) supports the generalizability of these criteria to the overall population of women aged 30 to 49 years in West Cameroon. However, the limited sample size and the fact that the study was conducted in a single center, do not allow to extend these results to the overall female population, especially considering the differences in HPV prevalence in other regions.

Strengths of the study are not discussed, and limitations are not fully covered (e.g.: sample size) and those that were mentioned are not widely explained; for instance, implications of having included a single centre.

Please see the completed paragraph above (lines 320-325) regarding limitations. The paragraph on strengths of the study was also completed as follows:

Lines 366-371: Strengths of our study included the application of ABCD criteria upon VIA examination in real-life conditions with immediate treatment when necessary, therefore supporting the feasibility of this “screen-and-treat” strategy. Furthermore, because all HPV-positive women underwent biopsy and cervical brushing regardless of the ABCD criteria results, there was no risk of verification bias in the calculations of sensitivity and specificity for ABCD criteria.

Implications of enhancing VIA with digital images should be more elaborated. Also, the feasibility of implementing this approach.

As mentioned in our response to one of the comments above, we are currently investigating the added value of digital images for VIA in a sub-sample of our study population. We agree that this topic requires further analysis and discussion, but data for the full sample presented in this article are not available to evaluate this issue.

Authors mention their 3T-Approach in the discussion, but it was barely mentioned in Methods.

Thank you for highlighting this. The following has been added to the Methods section:

Lines 96-99: This prospective study was carried out between September 2018 and March 2020 in the health district of Dschang (West Cameroon) as part of a 5-year cervical cancer screening programme. The screening strategy consisted of the “3T-Approach”, in which primary HPV Testing, Triage with VIA and Treatment are provided within one visit.

VERSION 2 – REVIEW

REVIEWER	Baena, Armando International Agency for Research on Cancer
REVIEW RETURNED	26-Dec-2021

GENERAL COMMENTS	Introduction The concept 3T-Approach is used in Methods and it sounds out of context if it is not previously introduced. I suggest authors briefly mention the 3T-Approach concept in the Introduction when they speak about triage. Methods

	The 3T-Approach concept sounds out of context. Please see suggestion above. Line 139: authors mentioned that the HPV test was given in one hour. Is this real for all participants? Seems to be perfect. Would you please better present summary measures such as mean or median with their respective variability (standard deviation or interquartile range) or a percentage of how many were done in less than 1 hour? If it is not possible to present a summary measure or percentage, I suggest authors word the sentence to avoid expectations that 100% of GeneXpert HPV results are available in one hour. Line 142: were healthcare providers mostly nurses? Were physicians considered part of this group? Sentences between lines 196 and 200 are confusing. Line 211: which version of Bethesda? 2015? ECC was first declared to have been endocervical curettage (lines 66 and 205) but later it is said to have been carried out with a brush (line 217). Please clarify. Line 235: is the public being currently kept informed even after the study? Statistical analyses: I think you also included the referral rate as a key measure of triage tests. Lines 270-269: if the study population included 340 participants (also stated in Figure 2), why does Table 1 describe 358? It seems that it should be 340 and this number should be consistent across the paper. Line 309: what this finding means? May this suggest that BCD criteria may improve the performance of VIA in women with TZ3 instead of A criteria alone? Normally, when TZ is type 3, VIA is not evaluable which may have implications for immediate treatment (thermal ablation of cryotherapy). This result deserves to be very well discussed but I was not able to find it in the discussion section. Line 310: were pathologists awarded VIA results and/or HPV results when interpreting cytology? Were they more likely to diagnose a positive result because of this information? This is also important to be mentioned when performance of cytology is analysed; even more when it is compared with other triage methods (Table 3). One of the main results of this paper is that Cytology has better performance (higher sensitivity and lower referral rate) compared to VIA and HPV partial genotyping. Therefore, conditions under cytology was performed should be mentioned/discussed. What were the advantages of cytology compared to VIA: personnel involved in cytology had more information than VIA examiners? ABCD criteria improved cytology? Etc. Line 311: this result based on histology (which is not valid because histology is part of the outcome) is not applicable (should not be mentioned anymore as authors dropped this invalid analysis from the main results). Lines 312-313: what about being a farmer? This is part of the socio-demographic characteristics and it was significant. The sentence needs revision. Can you please clarify what do you mean by “(CIN threshold not applicable)” in line 317? Line 341: “is not satisfactory” instead of “was not satisfactory” Lines 355-358: that’s true, which also affects the referral rate; lowering the threshold of positivity of VIA definitely improves the sensitivity but also increases the referral which is not desirable for a triage test.

	Lines 371-398: authors only discussed results of VIA vs partial genotyping; cytology is not discussed despite it seems to perform better. One may intuit that the main discussion between VIA and HPV genotyping is because of the possibility of offering immediate treatment (i.e., at the same screening visit). However, it is not that clear. Therefore, I suggest the authors elaborate a little bit more on this idea in terms of immediate treatment and the advantages and disadvantages of each screen-triage-triage approach. Given that the PPV for CIN2+ and especially for CIN3+ of HPV genotyping is higher than the PPV of VIA, and that the idea is to offer treatment immediately, I suggest authors explore/discuss the combined strategy of VIA and HPV genotyping which is currently recommended in 2021 WHO guidelines. I suggest this should be widely reported in both Results and Discussion. Also, it is important to mention how the enhanced VIA with digital images affected the waiting time and how feasible it would be to implement in a real-life scenario. In general, I consider that the Discussion still needs more elaboration. Line 399: it is not easy to follow the sentence since the first limitation is not easy to be identified. Lines 415-417: this is an overall strength of the study that also benefits the performance of cytology and HPV genotyping (not only ABCD criteria).

VERSION 2 – AUTHOR RESPONSE

Reviewer: 3

Dr. Armando Baena, International Agency for Research on Cancer

Comments to the Author:

Introduction

The concept 3T-Approach is used in Methods and it sounds out of context if it is not previously introduced. I suggest authors briefly mention the 3T-Approach concept in the Introduction when they speak about triage.

Thank you for this suggestion. The following has been added in the introduction:

Lines 70-73: A ‘3T-Approach’ (Test, Triage, Treat) combining testing with a rapid HPV test, triage of HPV-positive women with VIA, and treatment by thermal ablation of VIA-positive patients within the same day, has been previously used to further reduce the risk of loss to follow-up.

Methods

The 3T-Approach concept sounds out of context. Please see suggestion above.

Please refer to the response above.

Line 139: authors mentioned that the HPV test was given in one hour. Is this real for all participants? Seems to be perfect. Would you please better present summary measures such as mean or median with their respective variability (standard deviation or interquartile range) or a percentage of how many were done in less than 1 hour? If it is not possible to present a summary measure or percentage, I suggest authors word the sentence to avoid expectations that 100% of GeneXpert HPV results are available in one hour.

Detailed statistics for the time to obtain GeneXpert HPV results are not available for this study, but the Xpert HPV time to result has been documented in the "WHO Prequalification of In Vitro Diagnostics, WHO Reference Number: PQDx 0268-070-00" as approximately 60 minutes and the manufacurer Cepheid (www.cepheid.com) 56 minutes. This sentence has been modified as follows:

Lines 109-110: (…) participants undertook an HPV self-test (Self-HPV) that was subsequently analyzed by a point-of-care assay (GeneXpert®), with most results available within an hour.

Line 142: were healthcare providers mostly nurses? Were physicians considered part of this group?

Thank you for this comment. This has now been specified in the Methods.

Lines 113-116: Trained midwives performed gynecologic examination with VIA/VILI, assessment of ABCD criteria and transformation zone (TZ) type, and determined treatment modalities in a single visit. Two gynaecologists were available on call for a second opinion or advice.

Sentences between lines 196 and 200 are confusing.

We have modified this sentence for more clarity.

Lines 160-164: Women with positive ABCD criteria were eligible for treatment by thermal ablation, with the exception of (i) lesions extending into the endocervix which could not be covered by the probe tip, and (ii) suspicions of carcinoma, in-situ adenocarcinoma or invasive adenocarcinoma, which were referred to a gynaecologist to determine the need for further treatment (LLETZ or oncological management).

Line 211: which version of Bethesda? 2015?

The 2014 version of Bethesda was used. This has been added to the manuscript.

ECC was first declared to have been endocervical curettage (lines 66 and 205) but later it is said to have been carried out with a brush (line 217). Please clarify.

Thank you for pointing this out. Endocervical brushing was performed, which was analyzed by histology. This has been corrected throughout the text.

Line 235: is the public being currently kept informed even after the study?

Yes, as stated in the ‘Patient and public involvement’ section, a newsletter is published and disseminated among the community twice a year. This is ongoing and will be pursued until the end of the 3T study, which we have now specified in the manuscript.

Statistical analyses: I think you also included the referral rate as a key measure of triage tests.

Thank you for this comment. The positivity rate of triage tests has been added in the ‘Statistical analysis’ section.

Lines 270-269: if the study population included 340 participants (also stated in Figure 2), why does Table 1 describe 358? It seems that it should be 340 and this number should be consistent across the paper.

358 is the total number of participants for whom VIA was performed (with assessment by the ABCD criteria), while 340 is the number of women for whom valid histological results were available. This is explained on lines 231-234:

“Overall, 1964 women performed Self-HPV, of whom 361 (18·5%) had an HPV-positive test and underwent pelvic examination, three were excluded from the results analysis for lack of ABCD criteria assessment, and 340 (94·2%) had interpretable histology findings and constituted the study population”, and in Figure 2 (flowchart of participants). As Table 1 describes the study population according to ABCD criteria, all participants with available interpretation of ABCD criteria were included in this table.

Line 309: what this finding means? May this suggest that BCD criteria may improve the performance of VIA in women with TZ3 instead of A criteria alone? Normally, when TZ is type 3, VIA is not evaluable which may have implications for immediate treatment (thermal ablation of cryotherapy). This result deserves to be very well discussed but I was not able to find it in the discussion section.

Thank you for highlighting this. We have added a paragraph to discuss this result:

Lines 374-381: Having a TZ3 was associated with a better prediction of ABCD criteria compared to TZ1 (Table 2), which is unexpected as VIA is generally considered inadequate for the evaluation of TZ3 cervixes. This may be due to the use of B, C and D criteria in addition to acetowhiteness, enabling the detection of lesions extending to the ectocervix and bleeding in the absence of visible lesions. However, as A, B, C and D criteria were not assessed separately within this study sample, it is currently not possible to determine which criterion contributes most to a correct interpretation of VIA. A study is currently underway to assess each criterion individually for the detection of CIN2+.

Line 310: were pathologists awarded VIA results and/or HPV results when interpreting cytology? Were they more likely to diagnose a positive result because of this information? This is also important to be mentioned when performance of cytology is analysed; even more when it is compared with other triage methods (Table 3). One of the main results of this paper is that Cytology has better performance (higher sensitivity and lower referral rate) compared to VIA and HPV partial genotyping. Therefore, conditions under cytology was performed should be mentioned/discussed. What were the advantages of cytology compared to VIA: personnel involved in cytology had more information than VIA examiners? ABCD criteria improved cytology? Etc.

The conditions of cytological analysis have been specified in the Methods section, lines 175-177:

The cytotechnologists were aware of the HPV-positive status (but not of the HPV type) of participants but were blinded to the ABCD criteria interpretation.

Line 311: this result based on histology (which is not valid because histology is part of the outcome) is not applicable (should not be mentioned anymore as authors dropped this invalid analysis from the main results).

Thank you, this result has now been removed.

Lines 312-313: what about being a farmer? This is part of the socio-demographic characteristics and it was significant. The sentence needs revision.

The association between working as a farmer and correct prediction of ABCD criteria has been added in the Results section.

Lines 274-277: Overall, a correct prediction of the ABCD criteria was not impacted by the multiple sociodemographic characteristics of the population in the multivariate analysis, apart from women working as farmers who were less likely to have a correct prediction of ABCD criteria than employed women (OR 0.41, 95% CI 0.18-0.95).

Can you please clarify what do you mean by “(CIN threshold not applicable)” in line 317?

The positivity rate for each triage test was calculated for all HPV-positive participants, not only CIN2+ cases. The table has been modified for more clarity.

Line 341: “is not satisfactory” instead of “was not satisfactory”

Thank you, the change has been made.

Lines 355-358: that’s true, which also affects the referral rate; lowering the threshold of positivity of VIA definitely improves the sensitivity but also increases the referral which is not desirable for a triage test.

We agree with this comment, although we cannot strictly speak of “referral” in this particular setting as thermal ablation is performed during the pelvic exam after VIA assessment (same visit) without a need for referral. However, the impact of the low specificity and PPV on overtreatment rates is discussed on lines 322-341:

The low specificity and PPV, leading to higher overtreatment rates, arise because we considered any whitening to be positive, meaning many benign conditions (metaplasia, inflammation or other benign cervical changes) could produce false-positive results for the ABCD criteria. […] Overall, 54·4% of normal histology results and 71·4% of CIN1 were considered ABCD criteria positive and consequently underwent unnecessary treatment. Thus, 85% (174 of 205) of women who screened positive were treated without CIN2+. However, when considering all women screened for CC, including HPV-negative, 174 were treated unnecessarily out of 1964 screened by Self-HPV, corresponding to an overall 8·9% overtreatment rate in the total population screened. Despite the low specificity, our 3T-Approach in a single visit may be acceptable in an LMIC context because it reduces cost and loss to follow-up, which are recognized barriers to effective cervical cancer screening.(11,27) […] Furthermore, treatment by thermal ablation is associated with very low risks of side effects and morbidity.(29) Therefore, treatment of a significant number of false-positive cases in this context may be considered an acceptable strategy for effective control of CC in an LMIC setting and may contribute to reaching the target of the WHO’s elimination initiative.(3,5)

Lines 371-398: authors only discussed results of VIA vs partial genotyping; cytology is not discussed despite it seems to perform better. One may intuit that the main discussion between VIA and HPV genotyping is because of the possibility of offering immediate treatment (i.e., at the same screening visit). However, it is not that clear. Therefore, I suggest the authors elaborate a little bit more on this idea in terms of immediate treatment and the advantages and disadvantages of each screen-triage-triage approach. Given that the PPV for CIN2+ and especially for CIN3+ of HPV genotyping is higher than the PPV of VIA, and that the idea is to offer treatment immediately, I suggest authors explore/discuss the combined strategy of VIA and HPV genotyping which is currently recommended in 2021 WHO guidelines. I suggest this should be widely reported in both Results and Discussion. Also, it is important to mention how the enhanced VIA with digital images affected the waiting time and how feasible it would be to implement in a real-life scenario. In general, I consider that the Discussion still needs more elaboration.

Thank you for this suggestion; we have completed the discussion accordingly.

Lines 355-373: One of the screening strategies currently recommended by the WHO is combined HPV 16/18/45 genotyping (treated immediately) and VIA triage of non-16/18/45 HPV genotypes. In our study population, this combined strategy resulted in an increased sensitivity of 85.0%, but even further decreased the specificity and PPV, which would therefore even further increase overtreatment rates. On the contrary, triage by cytology (using a threshold of ASC-US for a positive triage) improved both sensitivity (80·0%, 95% CI 64·0-89·9) and specificity (87·5, 95% CI 83·1-90·7) compared to the ABCD criteria. However, although this strategy may be adapted to higher-middle and high-income countries, the lack of trained cytotechnicians and well-equipped laboratories in low-income countries, the higher cost, and the inability to provide same-day treatment to patients positively triaged with cytology, render this triaging strategy unsuitable for low-resource settings. In comparison, the ABCD criteria require only basic equipment at a low cost, and allow initiation of therapy without delay. […]

While digital imaging by smartphone may facilitate ABCD interpretation and enhance diagnostic performance, it may result in slightly prolonged examination time and may not be accessible in all settings.

Line 399: it is not easy to follow the sentence since the first limitation is not easy to be identified.

The discussion has been rearranged to improve the flow and structure.

Lines 415-417: this is an overall strength of the study that also benefits the performance of cytology and HPV genotyping (not only ABCD criteria).

We agree and have modified the sentence accordingly.

Lines 398-401: Furthermore, because all HPV-positive women underwent biopsy and cervical brushing regardless of the ABCD criteria results, there was no risk of verification bias in the calculations of sensitivity and specificity for all diagnostic strategies assessed.

VERSION 3 – REVIEW

REVIEWER	Baena, Armando International Agency for Research on Cancer
REVIEW RETURNED	05-Mar-2022

GENERAL COMMENTS	I thank the authors to have clarified some issues that I had before. I just want to make the last suggestion and it is to modify "relaxed IARC criteria" since it may cause confusion. Instead, I suggest authors say just "IARC criteria". Thank you, Armando